



STINFO COPY

AFRL-HE-WP-TP-2007-0003

**Object Recognition via Information-Theoretic
Measures/Metrics**

Daniel W. Repperger

Alan R. Pinkus

**Warfighter Interface Division
Battlespace Visualization Branch
Wright-Patterson AFB OH 45433-7022**

Julie A. Skipper

Christina D. Schrider

**Wright State University
Department of Biomedical
Human Factors and Industrial Engineering
Dayton OH 45435**

December 2006

Interim Report for the period November 2006 to December 2006

**Approved for public release;
Distribution is unlimited.**

**Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Interface Division
Battlespace Visualization Branch
Wright-Patterson AFB OH 45433**

| | | | | | |
|---|-------------------------|----------------------------------|--|--|---|
| REPORT DOCUMENTATION PAGE | | | | Form Approved OMB No. 0704-0188 | |
| <small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.</small> | | | | | |
| PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS. | | | | | |
| 1. REPORT DATE (DD-MM-YYYY) 1 Dec 2006 | | 2. REPORT TYPE Interim | | 3. DATES COVERED (From - To) November 2006 – December 2006 | |
| 4. TITLE AND SUBTITLE Object Recognition via Information-Theoretic Measures/Metrics | | | | 5a. CONTRACT NUMBER In-House | |
| | | | | 5b. GRANT NUMBER | |
| | | | | 5c. PROGRAM ELEMENT NUMBER 61102F | |
| 6. AUTHOR(S) Daniel W. Repperger* Alan R. Pinkus* Julie A. Skipper** Christina D. Schrider** | | | | 5d. PROJECT NUMBER 2313 | |
| | | | | 5e. TASK NUMBER HC | |
| | | | | 5f. WORK UNIT NUMBER 54 | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Wright State University Department of Biomedical** Human Factors and Industrial Engineering Dayton OH 45435 | | | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Materiel Command* Air Force Research Laboratory Human Effectiveness Directorate Warfighter Interface Division Battlespace Visualization Branch Wright Patterson AFB OH 45433-7022 | | | | 10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/HECV | |
| | | | | 11. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-HE-WP-TP-2007-0003 | |
| 12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited. | | | | | |
| 13. SUPPLEMENTARY NOTES AFRL/PA Cleared 1/12/2007, AFRL/WS-07-0076. | | | | | |
| 14. ABSTRACT Discrimination of friendly or hostile objects is investigated using information-theory measures/metric in an image which has been compromised by a number of factors. In aerial military images, objects with different orientation can be reasonably approximated by a single identification signature consisting of the average histogram of the object under rotations. Three different information-theoretic measures/metrics are studied as possible criteria to help classify the objects. | | | | | |
| 15. SUBJECT TERMS | | | | | |
| 16. SECURITY CLASSIFICATION OF: Unclassified | | | 17. LIMITATION OF ABSTRACT SAR | 18. NUMBER OF PAGES 14 | 19a. NAME OF RESPONSIBLE PERSON Daniel W. Repperger |
| a. REPORT U | b. ABSTRACT U | c. THIS PAGE U | | | 19b. TELEPHONE NUMBER (Include area code) |

Object recognition via information-theoretic measures/metrics

Daniel W. Repperger^{*1}, Alan R. Pinkus¹, Julie A. Skipper², Christina D. Schrider²

¹Air Force Research Laboratory, AFRL/HEC, Wright-Patterson AFB, Ohio, USA 45433;

²Wright State Univ., Dept. of Biomedical, Human Factors and Industrial Eng., Dayton, Ohio 45435

ABSTRACT

Discrimination of friendly or hostile objects is investigated using information-theory measures/metric in an image which has been compromised by a number of factors. In aerial military images, objects with different orientations can be reasonably approximated by a single identification signature consisting of the average histogram of the object under rotations. Three different information-theoretic measures/metrics are studied as possible criteria to help classify the objects. The first measure is the standard mutual information (MI) between the sampled object and the library object signatures. A second measure is based on information efficiency, which differs from MI. Finally an information distance *metric* is employed which determines the distance, in an information sense, between the sampled object and the library object. It is shown that the three (parsimonious) information-theoretic variables introduced here form an independent basis in the sense that any variable in the information channel can be uniquely expressed in terms of the three parameters introduced here. The methodology discussed is tested on a sample set of standardized images to evaluate their efficacy. A performance standardization methodology is presented which is based on manipulation of contrast, brightness, and size attributes of the sample objects of interest.

Keywords: Object discrimination, information theory, mutual information, parsimonious information measures, metrics

1. INTRODUCTION

The problem addressed in this paper is classical and rooted in the field of object detection and classification of objects. In this study, the emphasis, however, is on information-theoretic measures and a metric (Ref. 1, 3-5) to examine their possible efficacy in helping to discriminate an object in a visual image. This study will generalize a problem posed on improving sensitivity of discrimination (Ref. 2, 6) of objects but will employ information-theoretic criteria to better understand how the discrimination should be conducted. The basic problem of interest is first introduced.

1.1 The basic problem of interest

Fig. 1a portrays object 1 (friendly object) and Fig. 1b displays object 2 (hostile object) as a library template. The goal is to distinguish these objects in an image which may be compromised by a variety of factors. Some of the ways to compromise the object involve reducing its size, mitigating contrast, and altering the brightness to extremely bright or dark values. As a first step in this analysis, a comparison of the respective intensity histograms of the two original library objects is displayed in Fig. 2. In this paper the focus will be on the discrimination of the two objects using only a parsimonious number of variables of the information-theoretic type. It is noted that the sample images displayed in Figs. 1a-b were taken off web-based pictures freely available to the public and do not constitute any priority or military-specific information. The identification process proceeds in six steps;

Step 1: A test image is scanned for a possible friendly or hostile object.

Step 2: A sample from the test image is compared to possible template images of friendly or hostile objects.

Step 3: A distance norm, based on information-theoretic variables, determines a relative separation between the sample and each library object. Both information measures and a metric are considered in the norm definition.

* d.repperger@ieee.org; phone 1-937-255-8765; fax 1-937-255-8752

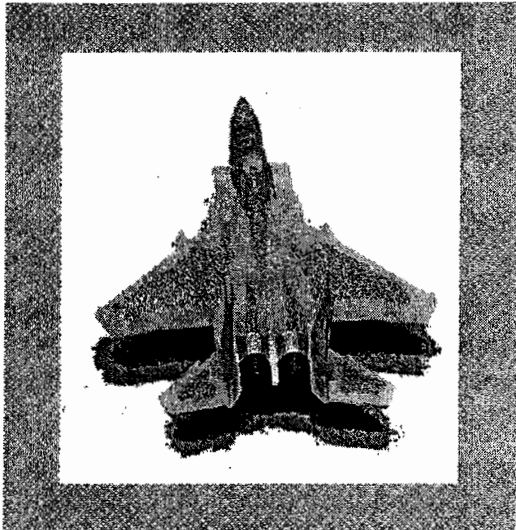
Step 4: A vote on the object's identity is based on the closeness of the sample to a particular library template.

Step 5: A voting scheme is then developed based on outputs of the constituent voters.

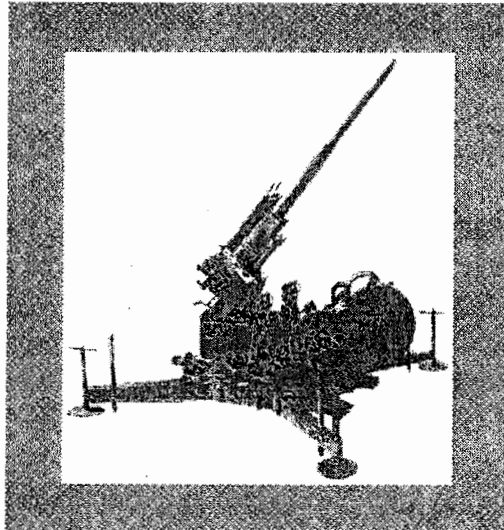
Step 6: The overall decision may depend in a nonlinear manner on the voting scheme (perhaps not majority).

A standard to fairly compare the efficacy of different object identification algorithms in images is presented next.

Good Object



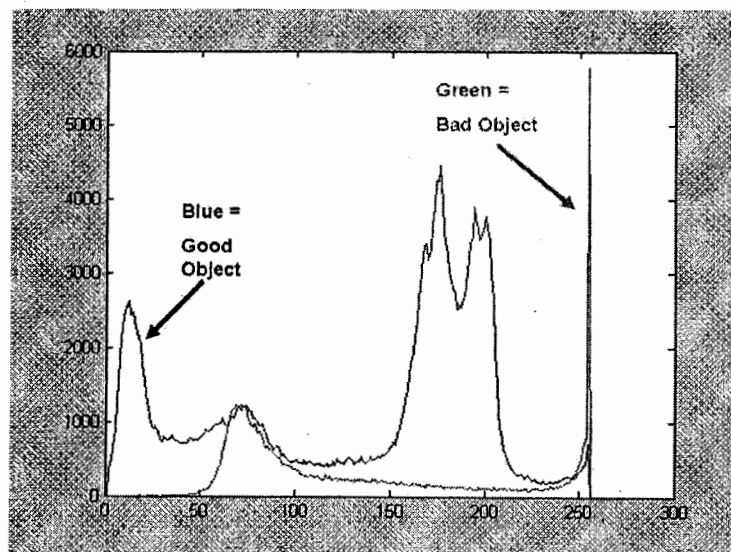
Bad Object



The Good Object versus the Bad Object for the Identification Problem of Interest

Fig. 1a – The good object (F-15A aircraft)

Fig. 1b – The bad object (anti-aircraft gun)



Histogram Signatures of the Good Object versus the Bad Object.

Fig. 2 – Histogram signatures of good versus bad object from Figs. 1a-b

2. A STANDARD TEST PROCEDURE FOR OBJECT RECOGNITION IN IMAGES

Adapting a concept discussed in Ref. 2 to provide a fair comparison on the efficacy of algorithms to detect objects, Fig. 3 portrays this notion. The goal is to display a possible means of objectively quantifying the ability of different

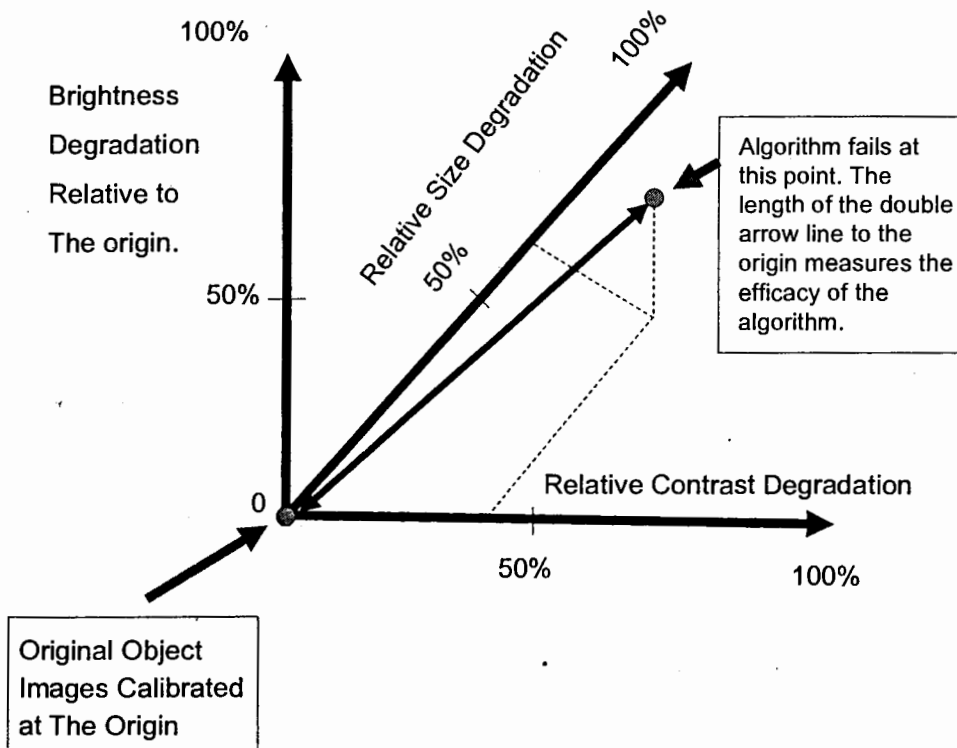


Fig. 3 Standard to compare object identification algorithms in images

algorithms to correctly discern between two objects, such as in Figs. 1a-b. Three major factors considered in this paper that influence the ability to discern objects in images include image contrast, brightness (amount of light being transmitted back to the observer from the image), and relative size of the key objects. The good and bad objects in Figs. 1a-b are at normal size, contrast, and intensity appear at the origin in Fig. 3 as an initial calibration. The three axes show various levels of degradation away from the origin in the directions of decreasing contrast, changing brightness, and reduction in size. The point in which an algorithm fails (e.g. less than 60% correct detection in a binary detection task) may define the limit of performance of the algorithm. Thus the distance length from the failure point to the origin in Fig. 3 is a possible measure to objectively state the efficacy of an object identification algorithm. Hence algorithms can be compared one against the other for their relative efficacy.

A brief discussion on basic information-theoretic concepts is presented as it pertains to the problem of correctly detecting objects in images.

3. PRELIMINARY INFORMATION THEORY CONSTRUCTS

3.1 Basic Definitions of the information channel variables

Fig. 4 represents an information-theoretic rendering of a channel (Ref. 3). Fig. 5, adapted from Ref. 1, portrays a

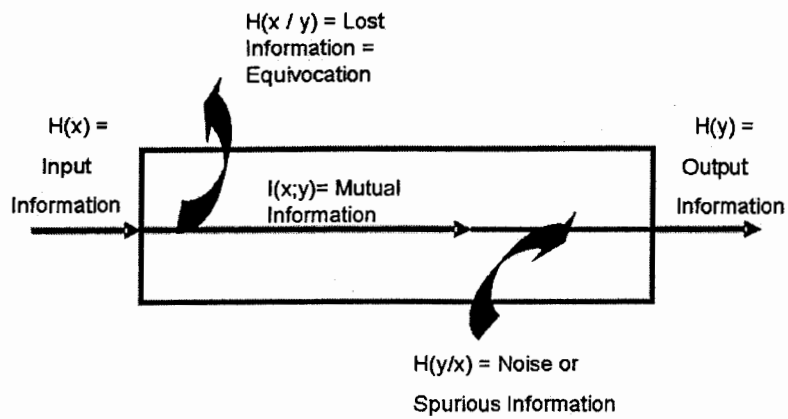


Fig. 4 - Basic elements of the information channel from Ref. 3

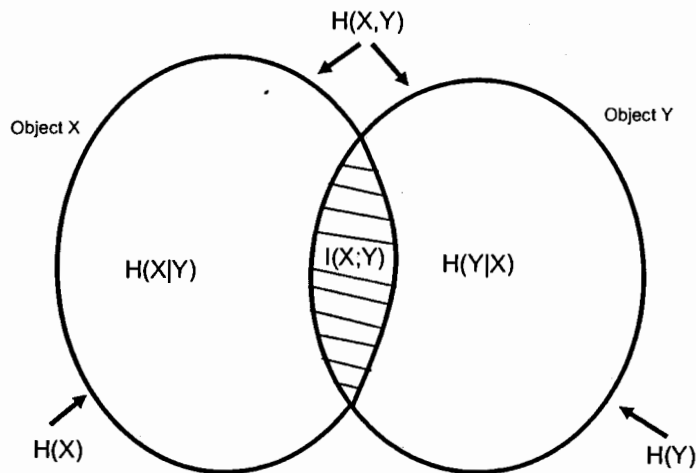


Fig. 5 – Venn diagram (Ref. 1) to illustrate the different entropies in an information channel

representation in terms of the various entropies in a Venn diagram. From Ref. 3, the five basic entities of an information channel, as originally defined by Shannon, are now presented. Let:

$H(x)$ = The input uncertainty in the input symbol set to the channel. (1)

$H(y)$ = The final output uncertainty of the output set to the channel. (2)

$H(x/y)$ = The equivocation lost to the environment. (3)

$$H(y/x) = \text{The spurious uncertainty provided by the environment on the channel.} \quad (4)$$

$$I(x;y) = \text{The mutual information transmitted through the information channel.} \quad (5)$$

Fig. 6 provides another rendering in which the mutual information variable $I(x;y)$ is considered within the context of the reduction of uncertainties.

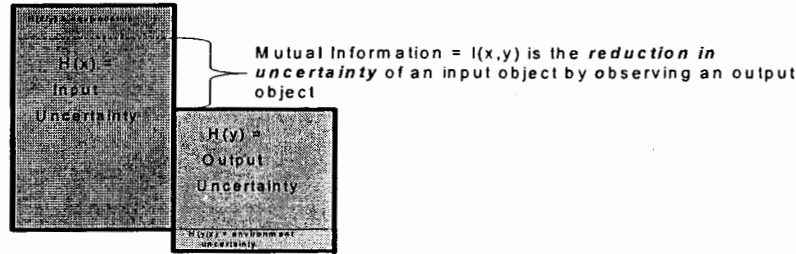


Fig. 6 – Definition of mutual information as reduction of uncertainties.

3.2 The basic definitions of the key variables in equations (1-5)

With more specificity, the details of equations (1-5) are now described. Let $p(.)$ represent the probability of an event. For an information channel with input symbol set $x \in X$, of size n , and received symbols $y \in Y$ at the output set of size q (q may not equal n), the following entropy ($H(.)$) relationships can be stated:

$$H(x) = \sum_{i=1}^n p(x_i) \log_2(1/p(x_i)) \quad (6)$$

$$H(y) = \sum_{j=1}^q p(y_j) \log_2(1/p(y_j)) \quad (7)$$

$$H(x,y) = \sum_{i,j}^{n,q} p(x_i, y_j) \log_2(1/p(x_i, y_j)) \quad (8)$$

$$H(x/y) = \sum_{i,j}^{n,q} p(x_i, y_j) \log_2(1/p(x_i | y_j)) \quad (9)$$

and

$$H(y/x) = \sum_{i,j}^{n,q} p(x_i, y_j) \log_2(1/p(y_j | x_i)) \quad (10)$$

In calculating all the uncertainty terms $H(.)$, if $p(.) = 0$ the contribution to the $H(.)$ variable is set to zero. Actually it can be shown, in a rigorous sense, that $\lim_{x \rightarrow 0} \{x \log(1/x)\} = \lim_{x \rightarrow 0} \{-x \log(x)\} \rightarrow 0$ so the contribution of a zero probability term to the $H(.)$ variable is, without question, zero.

3.3 The pertinent mathematical relationships between the key variables

A summary compendium of a number of important properties of the key variables (1-5) is now listed. From Figs. (4-6) and the basic definitions, equations (6-10), the following relationships can be shown to be true:

$$I(x;y) = H(x) + H(y) - H(x,y) \quad (11)$$

$$H(x/y) = H(x) - I(x;y) \quad (12)$$

$$H(y/x) = H(y) - I(x;y) \quad (13)$$

Since $H(x) > 0$, $H(y) \geq 0$, $H(x,y) \geq 0$, $H(x/y) \geq 0$, and $H(y/x) \geq 0$, it also follows that:

$$I(x;y) \geq 0 \quad (14)$$

$$I(x;y) = I(y;x) \quad (15)$$

$$I(x;x) = H(x) \quad (16)$$

$$I(x;y) \leq \min (H(x), H(y)) \leq H(x,y) \leq H(x) + H(y) \quad (17)$$

$$I(x;y) = H(x) - H(x/y) = H(y) - H(y/x) \quad (18)$$

3.4 Reformulation of the information theory problem in terms of parsimonious parameters

(The new variables make up an independent basis.)

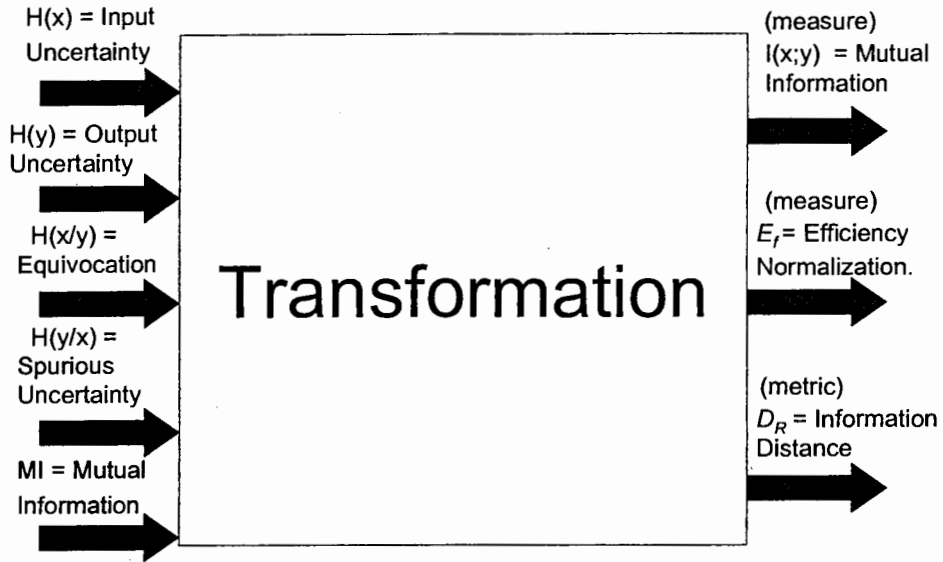


Fig 7 – A parsimonious redefinition of information theory variables.

In Fig. 7 a most parsimonious representation of the information channel is now presented. In this rendering, the four $H(.)$ quantities in equations (1-4) are considered only as uncertainty variables. The role of $(I(x;y))$ is to reduce uncertainty. Hence a relative information distance *metric* D_R is now defined as follows (Ref. 1):

$$D_R = H(x/y) + H(y/x) = H(x) + H(y) - 2 I(x;y) = 2 H(x,y) - H(x) - H(y) \quad (19)$$

and an efficiency measure E_f is introduced (Ref. 1):

$$E_f = (I(x;y) / H(x)), \quad \text{for } H(x) > 0. \quad (20)$$

Inspecting Fig. 7 it is seen on the left that the original information channel consists of five variables as defined by Shannon (Ref. 4) in which only three of them are independent. Also from Fig. 7, the three (parsimonious) variables D_R , E_f and $I(x;y)$ on the right are sufficient to completely define the information channel independently. Theorem 1 addresses this more parsimonious representation of the information channel.

Theorem 1: The three information variables D_R , E_f and $I(x;y)$ completely define the information channel, uniquely, using a bijective mapping in Fig. 7. All the five Shannon variables on the left side of the figure can be written in terms of these three key constituent information quantities (I , E_f and D_R) on the right side of Fig. 7 and vice versa.

The proof of Theorem 1 is given in Appendix A. For completeness, all eight relationships are listed here to show the unique, bijective, mapping that exists between the five uncertainty variables derived by Shannon and D_R , E_f and $I(x;y)$. The five Shannon variables satisfy:

$$H(x) = [I(x;y)] / E_f \quad \text{for } E_f > 0 \quad (21)$$

$$H(x/y) = [I(x;y) (1 - E_f)] / E_f \quad (22)$$

$$H(y/x) = D_R - I(x;y) (1 - E_f) / E_f \quad (23)$$

$$H(y) = I + D_R - I(x;y) (1 - E_f) / E_f \quad (24)$$

$$I(x;y) = I(x;y) \text{ (this variable was originally an information variable)} \quad (25)$$

Conversely, D_R , E_f and I on the right side of Fig. 7 satisfy:

$$D_R = H(x|y) + H(y|x) \quad (26)$$

$$E_f = [I(x;y)] / H(x), \quad \text{for } H(x) > 0 \quad (27)$$

$$I(x;y) = I(x;y) \quad (28)$$

Thus, there exists a unique, bijective, one-to-one mapping between the Shannon variables with the three parsimonious variables selected herein (D_R , E_f and I). Hence, only the variables D_R , E_f and $I(x;y)$ will be used in the sequel.

3.5 Physical properties of D_R and E_f :

First, it is important to discuss the physical interpretation of the two new introduced variables (D_R and E_f) considered in this paper. To better understand the utility of these variables, consider three cases: **Case 1:** The received symbols Y are independent of the input symbols X . **Case 2:** The received output symbols Y are precisely equal to the input symbols X (fully dependent or 100% correlated). The uncertainty terms $H(x/y)$ and $H(y/x)$ are both zero. **Case 3** will consider the intermediate situation where the received symbols Y are somehow related to the input symbols but the two uncertainty terms $H(x/y)$ and $H(y/x)$ may be non zero (positive). Table I shows a spectrum of dependence between input and output symbol sets in terms of all these three information-theoretic variables on the right side of Fig. 7.

Level of Dependency between the Input X and the Output Y

| | <div style="display: flex; justify-content: space-between; align-items: center;"> Low Dependency High Dependency </div> | | |
|------------|---|-------------------------------|-------------------------------|
| Variables | X and Y are Independent | X and Y are Related somewhat. | X and Y are highly Correlated |
| $H(x)$ | $H(x) = \text{same}$ | $H(x) = \text{same}$ | $H(x) = \text{same}$ |
| $H(y)$ | 0 | $0 < H(y) < H(x)$ | $H(y) = H(x)$ |
| $H(x/y)$ | High | Medium | 0 = Low |
| $H(y/x)$ | High | Medium | 0 = Low |
| $H(x,y)$ | High | Medium | Low = $H(x) = H(y)$ |
| $I(x,y)$ | 0 | $0 < I < H(x)$ | $I = H(x) = H(y)$ |
| $D_R(x,y)$ | High | Medium | 0 |
| $E_f(x,y)$ | 0 | $0 < E_f < 1$ | 1 |
| | Case 1 | Case 3 | Case 2 |

Table I – Range of values of the key variables in Fig. 7

3.6 Discussion of the entries in Table I

In Table I, the variable $H(x)$ is the input into the information channel, which is assumed to be invariant for this example. $H(y)$ can only get as large as $H(x)$ without interaction with the environmental term $H(y/x)$. The term $H(x/y)$ is the loss in bits from $H(x)$ to the environment, that are never recovered. $I(x,y)$ is the mutual information which represents a reduction in uncertainty that flows through the channel. $H(y/x)$ is the spurious entropy and $H(y)$ represents the received level of entropy at the channel's output.

The ranges of D_R and E_f are very interesting. For example, D_R is a relative information distance and when the random variables X and Y are independent (far apart from each other), D_R is at a maximum. However, when X and Y are 100% correlated, then D_R is zero. When X and Y fall between the extremes of being totally independent or totally correlated, then D_R is a positive number ($0 < D_R < 1$) indicating relative distance between the random variables. For the efficiency measure E_f , when the random variables X and Y are independent, then $I = 0$ and $E_f = 0$ indicating that the information channel is not efficient in producing information or reducing uncertainty. However, when X and Y are dependent, then $E_f = 1$, its largest value, so the information channel is maximally efficient in producing an information flow. For the intermediate case where X and Y have some correlation, then ($0 < E_f < 1$) and reflects the percent of information flowing in relation to its original input $H(x)$ and it is normalized, accordingly. It is noted that both $I(x,y)$ and E_f are measures, and not metrics but D_R is truly a relative information distance metric. We briefly discuss this distinction.

3.7 Measure versus metric properties of the information variables

It should be mentioned that $I(x,y)$ is a measure and not a metric. A true metric $\rho(x,y)$ must satisfy the following four relationships:

$$(M-1) \quad \rho(x,y) \geq 0 \text{ for all } x \text{ and } y. \quad (\text{positivity}) \quad (29)$$

$$(M-2) \quad \rho(x,y) = \rho(y,x) \quad (\text{similarity}) \quad (30)$$

$$(M-3) \quad \rho(x,z) \leq \rho(x,y) + \rho(y,z) \quad (\text{triangular inequality}) \quad (31)$$

$$(M-4) \quad \rho(x,x) = 0 \quad (32)$$

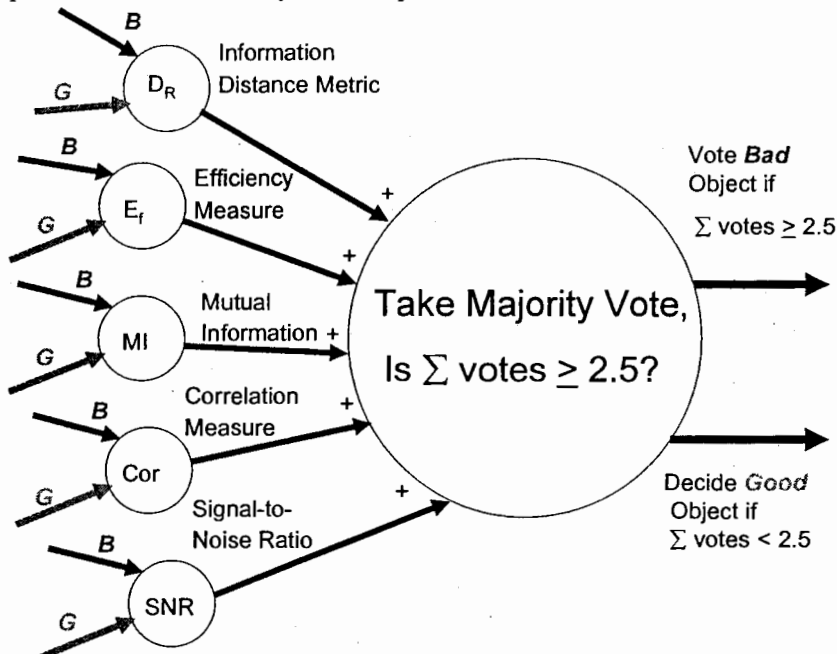
It is shown in Appendix B that $I(x;y)$ does not satisfy the triangular inequality, because, in general:

$$I(x;z) \leq I(x;y) + I(y;z) \quad (33)$$

does not hold for three random variables X, Y, and Z. Appendix B discusses the triangular inequality issues with respect to the measure variable $I(x;y)$ and the metric variable D_R . A study involving the three parsimonious information variables is conducted with the object images in Figs. 1a-1b and compared with the case of 5 variables. The other two variables include a standard correlation measure (sample object with each library object) and a signal to noise variable based on characteristics of the histograms of the sample to the library image.

4. APPLICATION TO THE OBJECT DISCRIMINATION PROBLEM

Fig. 8 displays a majority voting scheme in which up to five variables will be used to make a decision on whether the object in the image is either friendly or hostile. Case 1 will allow all five variables to make a majority vote. Case 2 will only consider the three parsimonious information variables. The decision criteria is that if the sum of the votes (Case 1) is greater or equal to 2.5, the choice 1 (bad object) is selected. The alternative is choice 0 (good object) if the sum of the votes is less than 2.5. For this test, the object in the image is corrupted with white-Gaussian noise and a signal detection theory approach is taken. A miss occurs when the object is a bad object, but the decision rule selects the good object. A false positive occurs for the case that the object selected is the bad object, when the ground truth is that the object is really the good object. Using this signal detection theory framework, the area under a ROC (receiver operator characteristic) curve is one method to evaluate performance. Using a technique analogous to Fig. 3, a Monte Carlo simulation was conducted. The level of noise intensity at which the ground truth figures become confused is a measure of the efficacy of the algorithm. For example, Fig. 9 shows the 5 majority voters for both objects. Fig. 10 shows a similar comparison for the use of only the three parsimonious information variables.



The Majority Voting Scheme for Improved Object Identification

Fig. 8 Majority voting scheme based on five constituent voters

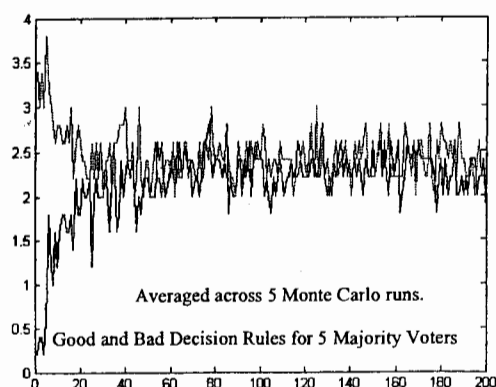


Figure 9 – Results of 5 Majority Voters

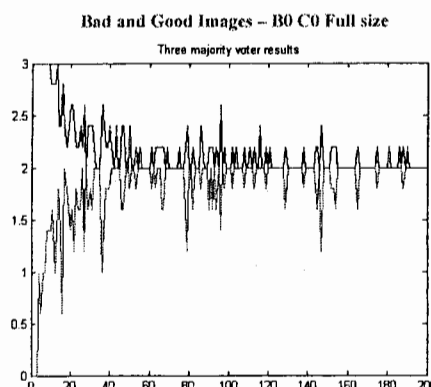


Figure 10 – Results of 3 Parsimonious variables

5. RESULTS

From Figs. 9 and 10, the intensity of the noise provides a relative comparison of the decision rules. As the noise power increased in this Monte Carlo simulation, the Fig. 10 decision making scheme performs better (larger noise strength before confusion occurs) for several reasons: (1) The voters represent an independent basis, (2) with fewer votes, the bias introduced by non independent voters is mitigated. In other words, little performance gain was achieved by adding two additional data streams of variables, if the new data stream was not pertinent to the decision making process.

6. CONCLUSIONS

The parsimonious variables selected herein seemed to show adequate performance in a majority voting scheme as compared to other standard measures used in the identification of objects in images. Both computational time and computational effort were saved when using the parsimonious set of information-theoretic variables as compared to a more complex simulation involving other variables.

REFERENCES

1. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., 1991.
2. D. W. Repperger, A. R. Pinkus, J. A. Skipper, and C. D. Schrider, "Stochastic Resonance Investigation of Object Detection in Images," *Proceedings of the IS&T / SPIE's XIX Annual Symposium on Electronic Imaging*, 2007, San Jose, California, USA.
3. T. B. Sheridan and W. R. Ferrell, *Man-Machine Systems: Information, Control, and Decision Models of Human Performance*, The MIT press, Cambridge, Massachusetts, 1981.
4. C. E. Shannon, "Communications in the Presence of Noise," *Proceedings of the IRE*, 37, pp. 10-22, 1949.
5. D. W. Repperger, M. S. Djouadi, et al., "Quantitative Measurements of System Complexity," USAF Patent application AFD 813, approved April, 2006, patent in progress.
6. D. W. Repperger, M. S. Djouadi, et. al., US patent #7,030,808 B1– "Nonlinear Target Recognition," 04/18/06.

APPENDIX A - RELATIONSHIPS BETWEEN ENTROPY AND INFORMATION VARIABLES

This appendix will demonstrate that equations (21-28) are valid. To show a one-to-one bijective mapping between the five Shannon uncertainty variables in equations (1-5) and the three information variables described in equations (18, 19 and 20), the five Shannon variables are first expressed uniquely terms of the three parsimonious quantities we wish to represent the channel (D_R , E_F and $I(x;y)$). It is noted that the mutual information term $I(x;y)$ appears both as a Shannon variable and as an information measure in this new formulation. Therefore it is only necessary to show the bijective mappings between D_R , E_F and the five Shannon variables: $I(x;y)$, $H(x)$, $H(y)$, $H(x/y)$, and $H(y/x)$. For notational

simplicity, let $D_R = x_1 = H(x/y) + H(y/x)$, $x_3 = I(x;y) = H(x) - H(x/y) = H(y) - H(y/x)$, and let $E_f = x_2 = I(x;y) / H(x) = x_3 / H(x)$. First, it can immediately be shown that $H(x) = x_3 / x_2$ (for $x_2 > 0$), so the first Shannon relationship is demonstrated. We now substitute this relationship for $H(x)$ into the expression: $x_3 = H(x) - H(x/y)$ to get the relationship $x_3 = (x_3)/(x_2) - H(x/y)$. Solving for $H(x/y)$ yields: $H(x/y) = (x_3) * (1-x_2)/(x_2)$. To get a similar relationship for $H(y/x)$ we use the symmetric property of the mutual information, i.e. $I(x;y) = H(x) - H(x/y) = H(y) - H(y/x) = x_3$. We find that $H(y/x) = x_1 - H(x/y) = x_1 - (x_3)(1-x_2)/(x_2)$. Finally knowing $H(y/x)$ we can reuse the mutual information relationship to recover $H(y)$, via $H(y) = x_3 + H(y/x)$ or $H(y) = x_3 + x_1 - x_3(1-x_2)/(x_2)$. Thus all five Shannon variables are now uniquely expressed in terms of the x_1 , x_2 , and x_3 variables selected here to provide a parsimonious representation of the information channel. Again, the practical presumptions were $H(x) > 0$ and $I > 0$, where it is assumed in equation (27) for $E_f > 0$. Conversely, to show the three information variables selected here (D_R , E_f and $I(x;y)$) can be similarly represented in terms of the five Shannon variables ($H(x)$, $H(y)$, $H(x/y)$, $H(y/x)$ and $I(x;y)$), it simply follows that $D_R = x_1 = H(x/y) + H(y/x)$ and $E_f = x_2 = x_3 / H(x) = I(x;y) / H(x)$ and finally $I(x;y) = I(x;y)$ to complete this demonstration.

APPENDIX B- METRIC AND MEASURE PROPERTIES OF D_R AND $I(X;Y)$

D_R enjoys the property of being a metric that satisfies the triangular inequality. $I(x;y)$, however, satisfies three of the four properties of a metric but does not satisfy the triangular inequality. This classifies $I(x;y)$ as a measure and not a metric. These points are demonstrated here for completeness.

B.1 Demonstration that D_R does satisfy all the properties of a metric

To show the D_R is a metric, Fig. B-1a is instructive using geometric arguments for two random variables X and Y (cf. Ref. 1). From Fig. B-1a, the quantities $H(x|y)$, $H(y|x)$ and $I(x;y)$ can now be specified in terms of the areas A_1 , A_2 , and A_3 as follows:

$$H(x|y) = A_1 \quad (B.1)$$

$$H(y|x) = A_3 \quad (B.2)$$

$$I(x;y) = A_2 \quad (B.3)$$

Fig. B-1b now generalizes this concept to three random variables X , Y , Z . In terms of the seven areas (A_1 - A_7) displayed, the following relationships become generalizations of Fig. B-1a into Fig. B-1b:

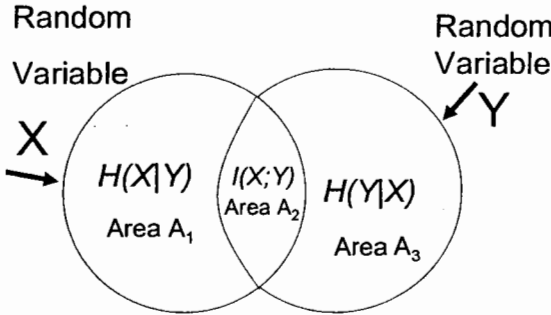


Figure B-1a Two Random Variables X and Y

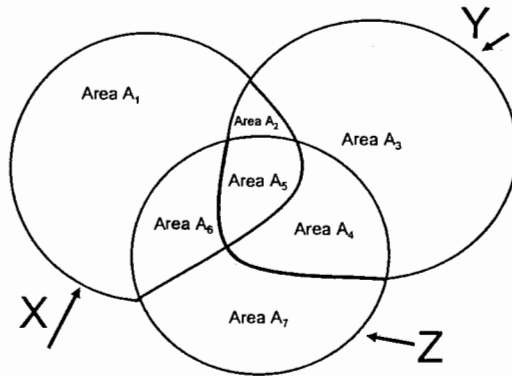


Figure B-1b – Three Random Variables X , Y , and Z

$$H(x|y) = A_1 + A_6,$$

$$H(y|x) = A_3 + A_4,$$

$$H(z|x) = A_4 + A_7,$$

$$H(x|z) = A_1 + A_2,$$

$$H(y|z) = A_2 + A_3,$$

$$H(z|y) = A_6 + A_7,$$

$$I(x;y) = A_2 + A_5$$

$$I(y;x) = A_5 + A_2$$

$$I(z;x) = A_5 + A_6$$

$$I(x;z) = A_6 + A_5$$

$$I(y;z) = A_5 + A_4$$

$$I(z;y) = A_4 + A_5$$

$$(B.4)$$

$$(B.5)$$

$$(B.6)$$

$$(B.7)$$

$$(B.8)$$

$$(B.9)$$

